Complete and Circularized Genome Assembly of a Human Isolate of *Vibrio navarrensis* Biotype *pommerensis* with MiSeq and MinION Sequence Data

Keike Schwartz,[a] Maria Borowiak,[a] Carlus Deneke,[a] Veronika Balau,[b] Claudia Metelmann,[b] ⓘ Eckhard Strauch[a]

aGerman Federal Institute for Risk Assessment, Department of Biological Safety, Berlin, Germany
bInstitut für Medizinische Diagnostik Labor Greifswald, Greifswald, Germany

ABSTRACT *Vibrio navarrensis* is a rare human pathogen. Strains of *Vibrio navarrensis* biotype *pommerensis* were isolated from seawater of the Baltic Sea. Recently, a strain of this biotype was recovered from a human patient. The isolate contains two circular chromosomes and a large plasmid with a size of 180 kb.

Vibrio navarrensis was found in river water and sewage in Spain (1). Later, isolates obtained from clinical samples suggested a human-pathogenic potential for this species (2). *V. navarrensis* biotype *pommerensis* strains were detected in seawater samples from the Baltic Sea (3). Isolates of the biotype *pommerensis* revealed major differences in biochemical profiles as well as borderline values in DNA-DNA hybridization experiments, compared to the *V. navarrensis* reference strain ATCC 51183 (4).

Strains of the biotype *pommerensis* have rarely been discovered in German coastal waters (5); however, in summer 2020, a biotype *pommerensis* isolate was recovered from a feverish patient suffering from an erysipelas-like infection on one leg. Prior to infection, the patient had been bathing in the Baltic Sea. The new isolate (20-VB00237) was subjected to whole-genome sequencing for further investigation.

The isolate was obtained from a blood culture (BD Bactec Plus Aerobic/Anaerobic-F) using Columbia agar with 5% sheep blood according to the manufacturer's recommendations (BD, Heidelberg, Germany). Subsequently, the isolate was cultivated for 24 h at 37°C in lysogeny broth. Genomic DNA was extracted using the PureLink genomic DNA minikit (Thermo Fisher Scientific, Waltham, MA, USA) and sequenced using MiSeq (Illumina, San Diego, CA, USA) and MinION (Oxford Nanopore Technologies [ONT], Oxford, UK) devices. An Illumina sequencing library was prepared using the Nextera DNA Flex kit. Paired-end sequencing was performed in 2×151-bp cycles on an Illumina MiSeq instrument using the MiSeq reagent kit v3 (600 cycles). Trimming of short reads using fastp v0.19.5 (6) resulted in 1.2 million high-quality reads (97.8% with a quality score of $\geq$Q30). An ONT sequencing library was prepared using the rapid barcoding kit and sequenced on an ONT MinION sequencer connected to an ONT MinIT v19.12.5 device (including Guppy base caller v3.2.10) using a FLO-MIN106 R9 flow cell. The reads obtained were trimmed using Porechop v0.2.3 (https://github.com/rrwick/Porechop), filtered using NanoFilt v2.7.1 (7), and quality checked using NanoStat v1.4.0 (7). In total, 25,772 reads with a read N_{50} value of 12,481 bp and a mean read quality score of 10.9 were available.

The two data sets were assembled and circularized using Unicycler v0.4.8 including Pilon (8–10). Default parameters were used for all software unless otherwise noted. The assembly resulted in two circular bacterial chromosomes, i.e., chromosome 1 (3,534,271 bp) and chromosome 2 (1,395,396 bp), and a closed plasmid (pVN20-VB00237) of 180,139 bp. The overall G+C content of the bacterial genome was
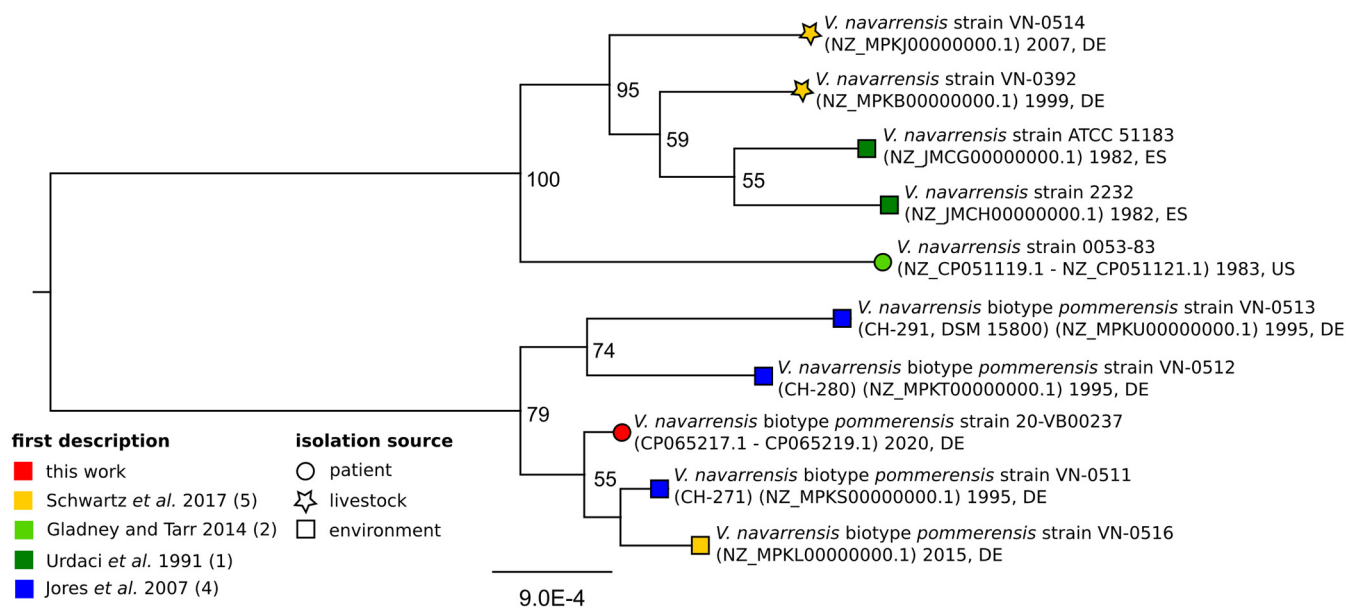
FIG 1 Best-scoring maximum likelihood tree based on a comparison of the amino acid sequences of 107 essential single-copy core genes of *V. navarrensis* biotype *pommerensis* strains and other published *V. navarrensis* strains using bcgTree. Numbers at the nodes designate bootstrap support values resulting from 100 bootstrap replicates. Numbers in parentheses are GenBank accession numbers. Dates in the node labels indicate the year of isolation, and abbreviations indicate the country of isolation (DE, Germany; ES, Spain; US, United States).

47.9%. The genome sequence was annotated using PGAP v4.11 (https://www.ncbi.nlm.nih.gov/genome/annotation_prok). Results revealed the presence of 4,662 coding sequences (CDSs) (4,511 protein-coding CDSs and 151 pseudogenes) and 150 RNA genes (34 rRNAs, 112 tRNAs, and 4 noncoding RNAs).

The isolate 20-VB00237 genome was compared to previously published *V. navarrensis* and *V. navarrensis* biotype *pommerensis* genomes. Therefore, all genomes were annotated using Prokka v1.1.3 (https://github.com/tseemann/prokka). Subsequently, phylogeny was inferred using bcgTree v1.1.0 (11), which compares the amino acid sequences of 107 single-copy core genes. The resulting maximum likelihood tree was visualized in Geneious v2020.2.2 (Biomatters, Auckland, New Zealand), manually rooted using the *V. navarrensis* node, and finalized using Inkscape. The final tree (Fig. 1) reveals that strain 20-VB00237 groups within a distantly related cluster of other *V. navarrensis* biotype *pommerensis* strains that were previously isolated from the Baltic Sea.

Data availability. The complete genome sequence of 20-VB00237 is available at NCBI (GenBank accession numbers CP065217 [chromosome 1], CP065218 [chromosome 2], and CP065219 [pVN20-VB00237]). Sequencing raw reads were deposited in the NCBI Sequence Read Archive (SRA) (accession numbers SRX9566546 [ONT data] and SRX9566545 [Illumina data]).

REFERENCES

1. Urdaci MC, Marchand M, Ageron E, Arcos JM, Sesma B, Grimont PA. 1991. *Vibrio navarrensis* sp. nov., a species from sewage. Int J Syst Bacteriol 41:290–294. https://doi.org/10.1099/00207713-41-2-290.
2. Gladney LM, Tarr CL. 2014. Molecular and phenotypic characterization of *Vibrio navarrensis* isolates associated with human illness. J Clin Microbiol 52:4070–4074. https://doi.org/10.1128/JCM.01544-14.
3. Jores J, Appel B, Lewin A. 2003. Cloning and molecular characterization of a unique hemolysin gene of *Vibrio pommerensis* sp. nov.: development of a DNA probe for the detection of the hemolysin gene and its use in identification of related *Vibrio* spp. from the Baltic Sea. FEMS Microbiol Lett 229:223–229. https://doi.org/10.1016/S0378-1097(03)00843-7.
4. Jores J, Appel B, Lewin A. 2007. *Vibrio navarrensis* biotype *pommerensis*: a new biotype of *V. navarrensis* isolated in the German Baltic Sea. Syst Appl Microbiol 30:27–30. https://doi.org/10.1016/j.syapm.2006.02.005.

5. Schwartz K, Kukuc C, Bier N, Taureck K, Hammerl JA, Strauch E. 2017. Diversity of *Vibrio navarrensis* revealed by genomic comparison: veterinary isolates are related to strains associated with human illness and sewage isolates while seawater strains are more distant. Front Microbiol 8:1717. https://doi.org/10.3389/fmicb.2017.01717.

6. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

7. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149.

8. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.

9. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. Microb Genom 3:e000132. https://doi.org/10.1099/mgen.0.000132.

10. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.

11. Ankenbrand MJ, Keller A. 2016. bcgTree: automated phylogenetic tree building from bacterial core genomes. Genome 59:783–791. https://doi.org/10.1139/gen-2015-0175.